PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : A61K 38/00, 15/31, 15/09, 48/00, C12N 15/79, 15/63, 15/00, C07K 16/00, C07H 21/00

A1

(11) International Publication Number: WO 97/18825

(43) International Publication Date: 29 May 1997 (29.05.97)

(21) International Application Number: PCT/US96/18370

(22) International Filing Date: 15 November 1996 (15.11.96)

(30) Priority Data:
60/006,882 17 November 1995 (17.11.95) US

(71) Applicant (*for all designated States except US*): THE UNIVERSITY OF BRITISH COLUMBIA [CA/CA]; 2075 Westbrook Mall, Vancouver, British Columbia V6T 1Z1 (CA).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): KALCHMAN, Michael [CA/CA]; #502-2233 Allisson Road, Vancouver, British Columbia V6T 1T7 (CA). HAYDEN, Michael, R. [US/CA]; 4484 West Seventh, Vancouver, British Columbia V6R 1W9 (CA).

(74) Agents: LARSON, Marina et al.; Oppedahl & Larson, Suite 309, 1992 Commerce Street, Yorktown Heights, NY 10598-4412 (US).

(81) Designated States: CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: PROTEIN WHICH INTERACTS WITH THE HUNTINGTON'S DISEASE GENE PRODUCT, cDNA CODING THEREFOR, AND ANTIBODIES THERETO

(57) Abstract

A protein, designated as HIP1, interacts differently with the gene product of a normal (16 CAG repeat) and an expanded (> 44 CAG repeat) HD gene. The HIP1 protein originally isolated from the yeast two-hybrid screen is encoded by a 1.2 kb cDNA, devoid of stop codons, that is expressed as a 400 amino acid polypeptide. By further screening of a human frontal cortex cDNA library, and employing the protocol for 5' Rapid Amplification of cDNA ends (RACE), a total of 4795 nucleotides (with an open reading frame of 914 amino acids) of the 10 kb message HIP1 have been isolated to date. Expression of the HIP1 protein was found to be limited to the brain, where the interaction of the HIP1 with the HD protein appears to be necessary for the association of the HD protein with the membrane or specific cytoskeletal components to render it functional. Because HIP1 interacts with expanded HD protein less well than with normal length HD, introduction of additional HIP1 or overexpression of HIP-1 can lead to increased functionality of the defective or normal HD protein. Alternatively, modified forms of the HIP1 which bind more effectively to expanded HD could be introduced to convert the expanded HD protein into a functional molecule.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgystan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic			SE	Sweden
CG	Congo	KR	Republic of Korea	SG	Singapore
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LR	Liberia	SZ	Swaziland
CS	Czechoslovakia	LT	Lithuania	TD	Chad
CZ	Czech Republic	LU	Luxembourg	TG	Togo
DE	Germany	LV	Latvia	TJ	Tajikistan
DK	Denmark	MC	Monaco	TT	Trinidad and Tobago
EE	Estonia	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	UG	Uganda
FI	Finland	ML	Mali	US	United States of America
FR	France	MN	Mongolia	UZ	Uzbekistan
GA	Gabon	MR	Mauritania	VN	Viet Nam

PROTEIN WHICH INTERACTS WITH THE HUNTINGTON'S DISEASE GENE PRODUCT, cDNA CODING THEREFOR, AND ANTIBODIES THERETO

BACKGROUND OF THE INVENTION

This application relates to a protein designated as HIP1 which interacts with the Huntington's Disease gene product, cDNA coding for HIP1, and methods and compositions relating thereto.

"Interacting proteins" are proteins which associate *in vivo* to form specific stable complexes. Non-covalent bonds, including hydrogen bonds, hydrophobic interactions and other molecular associations form between the proteins when two protein surfaces are matched or have affinity for each other. This affinity or match is required for the recognition of the two proteins, and the formation of a stable interaction. Protein-protein interactions are involved in the assembly of enzyme subunits; in antigen-antibody reactions; in forming the supramolecular structures of ribosomes, filaments, and viruses; in transport; and in the interaction of receptors on a cell with growth factors and hormones.

Huntington's disease is an adult onset disorder characterized by selective neuronal loss in discrete regions of the brain and spinal chord that lead to progressive movement disorder, personality change and intellectual decline. From onset, which generally occurs around age 40, the disease progresses with worsening symptoms, ending in death approximately 18 years after onset.

The biochemical cause of Huntington's disease has thus far not been determined. Various theories have been advanced, but each has failed to stand up to experimental evidence designed to test its validity. For example, it was suggested that the selective neuronal loss could be attributed to restricted expression of mRNA or proteins in cells undergoing degeneration. No obviously altered levels of mRNA transcript or protein expression has ever been observed in HD-affected tissues, however.

While the biochemical cause of Huntington's disease has remained elusive, a mutation in a gene within chromosome 4p16.3 subband has been identified and linked to the disease. This gene, referred to as the Huntington's Disease or HD gene, contains three repeat regions, a CAG repeat region and two CCG repeat regions. Testing of Huntington's disease patients has shown that the CAG region is highly polymorphic, and that the number of CAG

repeat units in the CAG repeat region is a very reliable diagnostic indicator of having inherited the gene for Huntington's disease. Thus, in control individuals and in individuals suffering from neuropsychiatric disorders other than Huntington's disease, the number of CAG repeats is between 9 and 35, while in individuals suffering from Huntington's disease the number of CAG repeats is expanded and is 36 or greater.

The protein product encoded by the HD gene has been localized to the cytoplasm, including to the membranes of vesicles on the brain of both normal and HD-affected individuals. To date, no differences have been observed at either the total RNA, mRNA or protein levels between normal and HD-affected individuals. Thus, the function of the HD protein and its role in the pathogenesis of Huntington's Disease remain to be elucidated.

SUMMARY OF THE INVENTION

We have now identified a protein, designated as HIP1, that interacts differently with the gene product of a normal (16 CAG repeat) and an expanded (>44 CAG repeat) HD gene. The HIP1 protein originally isolated from the yeast two-hybrid screen is encoded by a 1.2 kb cDNA, devoid of stop codons, that is expressed as a 400 amino acid polypeptide. By further screening of a human frontal cortex cDNA library, and employing the protocol for 5' Rapid Amplification of cDNA ends (RACE), a total of 4795 nucleotides (with an open reading frame of 914 amino acids) of the 10 kb message HIP1 have been isolated to date. Expression of the HIP1 protein was found to be limited to the brain, where the interaction of the HIP1 with the HD protein appears to be necessary for the association of the HD protein with the membrane or specific cytoskeletal components to render it functional. Because HIP1 interacts with expanded HD protein less well than with normal length HD, introduction of additional HIP1 or overexpression of HIP-1 can lead to increased functionality of the defective or normal HD protein. Alternatively, modified forms of the HIP1 which bind more effectively to expanded HD could be introduced to convert the expanded HD protein into a functional molecule.

BRIEF DESCRIPTION OF THE DRAWING

Fig. 1 graphically depicts the amount of interaction between HIP1 and Huntingtin proteins with varying lengths of polyglutamine repeat.

DETAILED DESCRIPTION OF THE INVENTION

The HIP1 protein which interacts with the HD gene product was identified using the yeast two-hybrid system described in US Patent No. 5,283,173 which is incorporated herein by reference. Briefly, this system utilizes two chimeric genes or plasmids expressible in a yeast host. The yeast host is selected to contain a detectable marker gene having a binding site for the DNA binding domain of a transcriptional activator. The first chimeric gene or plasmid encodes a DNA-binding domain which recognizes the binding site of the selectable marker gene and a test protein or protein fragment. The second chimeric gene or plasmid encodes for a second test protein and a transcriptional activation domain. The two chimeric genes or plasmids are introduced into the host cell and expressed, and the cells are cultivated. Expression of the detectable marker gene only occurs when the gene product of the first chimeric gene or plasmid binds to the DNA binding domain of the detectable marker gene, and a transcriptional activation domain is brought into sufficient proximity to the DNA-binding domain, an occurrence which is facilitated by protein-protein interactions between the first and second test proteins. By selecting for cells expressing the detectable marker gene, those cells which contain chimeric genes or plasmids for interacting proteins can be identified, and the gene can be recovered and identified.

In testing for Huntington Interacting Proteins, several different plasmids were prepared containing portions of the HD gene. The first four, identified as 16pGBT9, 44pGBT9, 80pGBT9 and 128pGBT9, were GAL4 DNA binding domain-HD in-frame fusions containing nucleotides 314 to 1955 (amino acids 1-540) of the published HD cDNA sequences cloned into the vector pGBT9 (Clontech). These plasmids contain a CAG repeat region of 16, 44, 80 and 128 glutamine-encoding repeats, respectively. A clone (DMK BamHIpGBT9) was made by fusing acDNA encoding the first 544 amino acids of the myotonic dystrophy gene (a gift from R. Komeluk) in-frame with the GAL4-DNA BD of pGBT9 and was used as a negative control.

These plasmids have been used to identify and characterize HIP1, two additional HD-interacting proteins, HIP2 and HIP3 proteins, and can be further used for the identification of additional interacting proteins, and for tests to refine the region on the protein in which the interaction occurs. Thus, a first aspect of the invention is these four plasmids, and the use of this plasmids in identifying HD-interacting proteins. Furthermore, it will be appreciated that the GAL4 DNA-binding and activating domains are not the only domains which can be used in the yeast two-hybrid assay. Thus, in a broader sense, the invention encompasses any chimeric genes or plasmids containing nucleotides 314 to 1955 of the HD gene together with an activating or DNA-binding domain suitable for use in the yeast one, two- or three-hybrid assay for proteins critical in either binding to the HD protein or responsible for regulated expression of the HD gene.

After introducing the plasmids into Y190 yeast host cells, transforming the host cells with an adult human brain Matchmaker™ (Clontech) cDNA library coupled with a GAL4 activating domain, and selecting for the expression of two detectable marker genes to identify clones containing genes for interacting proteins, the activating domain plasmids were recovered and analyzed. As a result of this analysis, three different cDNA fragments were identified as encoding for HD-interacting proteins and designated as HIP1, HIP2 and HIP3. The sequences of HIP1 and HIP3 are given in Seq. ID. Nos 1 and 3. The polypeptides which each encodes are given by Seq. ID Nos. 2 and 4. Further investigation of the HIP1 cDNA resulted in the characterization of an additional region of cDNA totaling 4795 bases and a corresponding protein, the sequences of which are given by Seq ID Nos. 5 and 6, respectively.

The cDNA molecules, particularly those encoding portions of HIP1, can be explored using oligonucleotide probes for example for amplification and sequencing. In addition, oligonucleotide probes complementary to the cDNA can be used as diagnostic probes to localize and quantify the presence of HIP1 DNA. Probes of this type with a one or two base mismatch can also be used in site-directed mutagenesis to introduce variations into the HIP1 sequence which may increase. Thus, a further aspect of the present invention is an oligonucleotide probe, preferably having a length of from 15-40 bases which specifically and selectively hybridizes with the cDNA given by Seq. ID No. 1 or 5 or a sequence complementary thereto. As used herein, the phrase "specifically and selectively hybridizes with" the

cDNA refers to primers which will hybridize with the cDNA under stringent hybridization conditions.

DNA sequencing of the HIP1 cDNA initially isolated from the yeast two-hybrid screen revealed a 1.2 kb cDNA that shows no significant degree of nucleic acid identity with any stretch of DNA using the blastn program at ncbi (blast@ncbi.n1m.nih.gov). When the entire HIP1 cDNA sequence (SEQ ID NO. 5) is translated into a polypeptide, the entire HIP1 cDNA coding (nucleotides 328-3069) is observed to be devoid of stop codons, and to produce a 914 amino acid polypeptide. A polypeptide identity search revealed an identity match over the entire length of the protein (46% conservation) with that of a hypothetical protein from *C. elegans* (ZK370.3 protein; *C. elegans* cosmid ZK370). This *C. elegans* protein shares identity with the mouse talin gene, which encodes a 217 kDa protein implicated with maintaining integrity of the cytoskeleton. It also shares identity with the SLA2/MOP2/ END4 gene from *Saccharomyces cerevisiae*, which is known to code for an essential cytoskeletal associated gene required for the accumulation and or maintenance of plasma membrane H^+-ATPase on the cell surface. When pairwise comparisons are performed between HIP1 and the C. elegans ZK370.3 protein (Genpept accession number celzk370.3), it shows 26% complete identity and an overall 46% level of conservation. Comparative analysis between HIP1 and SLA2/MOP2/ END4 (EMBL accession number Z22811) demonstrate similar conservation (20% identity, 40% conservation).

HIP2 is a 2.0 kb cDNA that encodes all but the 5'-most 33 amino acids of human $E2_{25k}$ ubiquitin conjugating enzyme. The resulting peptide has 100% identity with the previously characterized bovine $E2_{25k}$ protein. The cDNA has 95% nucleotide identity with the bovine cDNA. Ubiquitin-conjugating enzyme is an important component in ubiquitin-mediated protein degradation pathways.

No difference in the strength of the interaction between HIP2 and HD constructs containing either 44 or 15 CAG repeats is detected using a quantitative β-galactosidase assay. The expression pattern of HIP2 ($E2_{25k}$) in the various parts of the brain and nervous system appears to follow the specific neuropathology observed in HD, although there does not appear to be any difference in expression levels between HD-affected and HD-non-affected individuals.

The third cDNA encoding an HD-interacting protein is a 537 bp cDNA coding for 187 amino acids. A search of known DNA databases did not identify the sequence homology with any known genes. However, when a protein search was performed using the blatsp server, a strong identity between HIP3 and ankyrin-related proteins was observed. The strongest identity was with the D2021.8 gene product of *C. elegans*, an uncharacterized gene, but there is also a 41% identity with AKR1, a yeast ankyrin repeat-containing protein. Furthermore, when analogous structures with charge conservation over the same amino acid stretch are considered, there is 70% protein identity. HIP3 also shares approximately 60% amino acid conservation with human brain specific ankyrins (ankyrin B and ankyrin C). Thus, it is reasonable to conclude that HIP3, like known ankyrins, is a cytoskeletal protein, and may be involved, like previously characterized ankyrins in promoting interactions between the membrane skeleton and other membrane proteins.

Further exploration of these three HD interacting proteins revealed several important facts about HIP1 that implicate it in a significantly in the pathogenesis of Huntington's Disease. First, as shown in Fig. 1, it was found that the strength of the interaction between HD protein and HIP1 is dependent on the number of CAG repeats. Second, it was found that expression of the HIP1 protein is not ubiquitous, but is limited to brain tissue. The highest amounts of expression are in the cortex, with lower levels being seen in the cerebellum, caudate and putamen.

Both HIP1 and HIP3 appear to be proteins which are involved in the maintaining the structural integrity of the cytoskeleton and various components of the cellular membrane, including microtubules and focal adhesions. Based upon this, the HD protein may be associated as part of the cytoskeletal matrix in cells where it is expressed, and our work supports the conclusion that binding of HIP1 to the HD protein is necessary for the functional incorporation of the HD protein into the cell membrane. In this circumstance, the larger polyglutamine tract in huntingtin has a decreased ability for an HIP1-HD interaction. This decreased affinity for each other disrupts the normally strong HD-HIP1-cytoskeletal anchoring association. Further, the HIP1-HD interaction may be a critical interaction at the membranes of synaptic vesicles and a decrease in the affinity of HIP1 for huntingtin may affect protein trafficking or membrane organization throughout

the neuron. Finally, we have demonstrated that HIP1 and HD are both found in the Triton X-100 insoluble membrane compartment of the cell, therefore, a decreased interaction between HIP1 and huntingtin may allow an abnormally subtle amount of huntingtin to be found in subcellular compartments in which it is normally found.

As a result of all three of these phenomenon, increased apoptosis can occur in specific neurons within the striatum. This increase in apoptosis arises from an increased susceptibility of polyglutamine-expanded huntingtin to cleavage by apopain, and because more of the expanded forms of the HD protein may be available for cleavage (and subsequent apoptosis) due to the fact they are not as tightly associated at the HD-HIP1-cytoskeletal complex.

This understanding of a biochemical basis for the pathogenesis of Huntington's Disease opens the doorway to a therapeutic method to ameliorate the pathology in patients expressing huntingtin protein with expanded polyglutamine tracts. In accordance with the method, the patient is treated to increase the amount of HIP1 or an equivalent polypeptide which interacts less well with expanded Huntingtin than with Huntingtin having a CAG repeat region containing 15 to 35 repeats and facilitates the incorporation of Huntingtin into brain cell membranes.

Increasing expression of HIP1 or an equivalent polypeptide can be accomplished using gene therapy approaches. In general, this will involve introduction of DNA encoding HIP1 in an expressable vector into the brain cells. Vectors which have been shown to be suitable expression systems in mammalian cells include the herpes simplex viral based vectors: pHSV1 (Geller et al. Proc. Natl. Acad. Sci 87:8950-8954 (1990)); recombinant retroviral vectors: MFG (Jaffee et al. Cancer Res. 53:2221-2226 (1993)); Moloney-based retroviral vectors: LN, LNSX, LNCX, LXSN (Miller and Rosman Biotechniques 7:980-989 (1989)); vaccinia viral vector: MVA (Sutter and Moss Proc. Natl. Acad. Sci. 89:10847-10851 (1992)); recombinant adenovirus vectors : pJM17 (Ali et al Gene Therapy 1:367-384 (1994)), (Berkner K. L. Biotechniques 6:616-624 1988); second generation adenovirus vector: DE1/DE4 adenoviral vectors (Wang and Finer Nature Medicine 2:714-716 (1996)); and Adeno-associated viral vectors: AAV/Neo (Muro-Cacho et al. J. Immunotherapy 11:231-237 (1992)).

Delivery of retroviral vectors to brain and nervous system tissue has been described in US Patents Nos. 4,866,042, 5,082,670 and 5,529,774, which are incorporated herein by references. These patents disclose the use of cerebral grafts or implants as one mechanism for introducing vectors bearing therapeutic gene sequences into the brain, as well as an approach in which the vectors are transmitted across the blood brain barrier.

In addition to increasing the amount of HIP1 present in brain cells of affected individuals, HD lethal phenotype may be rescued by coexpression of a HIP1 and normal sized HD protein within the same cell, specifically within neurons. The over-expression of the normal HD protein and the presence of excess HIP1 in the cell may be able to override the damaging effects of a decreased interaction between HIP1 and an expanded form of the HD protein. Therefore, a "normal state" of interaction of HD with HIP1 will rescue the cell from premature apoptotic death. Thus, a therapeutically desirable mammalian expression vector may include both a region encoding HIP1 and a region encoding normal (less than 35 repeats) HD protein.

To further illustrate the methods of making the materials which are the subject of this invention, and the testing which has established their utility, the following non-limiting experimental procedures are provided.

EXAMPLE 1
IDENTIFICATION OF INTERACTING PROTEINS

GAL4-HD cDNA constructs

An HD cDNA construct (44pGBT9), with 44 CAG repeats was generated encompassing amino acids 1 - 540 of the published HD cDNA . This cDNA fragment was fused in frame to the GAL4 DNA-binding domain (BD) of the yeast two-hybrid vector pGBT9 (Clontech). Other HD cDNA constructs, 16pGBT9, 80pGBT9 and 128pGBT9 were constructed, identical to 44pGBT9 but included only 16, 80 or 128 CAG repeats, respectively.

Another clone (DMKDBamHIpGBT9) containing the first 544 amino acids of the myotonic dystrophy gene (a gift from R. Korneluk) was fused in-frame with the

GAL4-DNA BD of pGBT9 and was used as a negative control. Plasmids expressing the GAL4-BDRAD7 (D. Gietz, unpublished) and SIR3 were used as a positive control for the β-galactosidase filter assay.

The clones IT15-23Q, IT15-44Q and HAP1 were generous gifts from Dr. C. Ross. These clones represent a previously isolated huntingtin interacting protein that has a higher affinity for the expanded form of the HD protein.

Yeast strains, transformations and β-galactosidase assays

The yeast strain Y190 (MATa leu2-3,112, ura3-52, trp1-901, his3-Δ200, ade2-101, gal4Δgal80Δ, URA3::GAL-lacZ, LYS2::GAL-HIS3,cyc^r) was used for all transformations and assays. Yeast transformations were performed using a modified lithium acetate transformation protocol and grown at 30 C using appropriate synthetic complete (SC) dropout media.

The β-galactosidase chromogenic filter assays were performed by transferring the yeast colonies onto Whatman filters. The yeast cells were lysed by submerging the filters in liquid nitrogen for 15-20 seconds. Filters were allowed to dry at room temperature for at least five minutes and placed onto filter paper presoaked in Z-buffer (100 mM sodium phosphate (pH7.0) 10 mM KCl, 1 mM $MgSO_4$) supplemented with 50 mM 2-mercaptoethanol and 0.07 mg/ml 5-bromo-4-chloro-3-indolyl β-D-galactoside (X-gal). Filters were placed at 37 C for up to 8 hours.

Yeast two-hybrid screening for huntingtin interacting protein (HIP)

cDNAs from an human adult brain Matchmaker™ cDNA library (Clontech) was transformed into the yeast strain Y190 already harboring the 44pGBT9 construct. The transformants were plated onto one hundred 150 mm x 15 mm circular culture dishes containing SC media deficient in Trp, Leu and His. The herbicide 3-amino-triazole (3-AT) (25mM) was utilized to limit the number of false His+ positives (31). The yeast transformants were placed at 30 C for 5 days and β-galactosidase filter assays were performed on all colonies found after this time, as described above, to identify β-galactosidase+ clones. Primary His+/β-galactosidase+ clones were then orderly patched onto a grid on SC

-Trp/-Leu/-His (25 mM 3AT) plates and assayed again for His+ growth and the ability to turn blue with a filter assay. Secondary positives were identified for further analysis. Proteins encoded by positive cDNAs were designated as HIPs (Huntingtin Interactive Proteins). Approximately 4.0 x 10^7 Trp/Leu auxotrophic transformants were screened and of 14 clones isolated 12 represented the same cDNA (HIP1), and the other 2 cDNAs, HIP2 and HIP3 were each represented only once.

The HIP cDNA plasmids were isolated by growing the His+/β-galactosidase+ colony in SC -Leu media overnight, lysing the cells with acid-washed glass beads and electroporating the bacterial strain, KC8 (leuB auxotrophic) with the yeast lysate. The KC8 ampicillin resistant colonies were replica plated onto M9 (-Leu) plates. The plasmid DNA from M9+ colonies was transformed into DH5-a for further manipulation.

EXAMPLE 2

CONFIRMATION OF INTERACTIONS

The HIP1-GAL4-AD cDNA activated both the lac-Z and His reporter genes in the yeast strain Y190 only when co-transformed with the GAL4-BD-HD construct, but not the negative controls (Figure 1) of the vector alone or a random fusion protein of the myotonin kinase gene. In order to assess the influence of the polyglutamine tract on the interaction between HIP1 and HD, semi-quantitative β-galactosidase assays were performed. GAL4-BD-HD fusion proteins with 16, 44, 80 and 128 glutamine repeats were assayed for their strength of interaction with the GAL4-AD-HIP1 fusion protein.

Liquid β-galactosidase assays were performed by inoculating a single yeast colony into appropriate synthetic complete (SC) dropout media and grown to OD600 0.6-1.5. Five millilitres of overnight culture was pelleted and washed once with 1 ml of Z-Buffer, then resuspended in 100 ml Z-Buffer supplemented with 38 mM 2-mercapto-ethanol, and 0.05% SDS. Acid washed glass beads (~100 ml) were added to each sample and vortexed for four minutes, by repeatedly alternating a 30 seconds vortex, with 30 seconds on ice. Each sample was pelleted and 10 ml of lysate was added to 500 ml of lysis buffer. The samples were incubated in a 30 C waterbath for 30 seconds and then 100 ml of a 4 mg/ml o-nitrophenyl b-D galactopyranoside (ONPG) solution was added to each tube.

The reaction was allowed to continue for 20 minutes at 30 C and stopped by the addition of 500 ml of 1 M Na_2CO_3 and placing the samples on ice. Subsequently, OD420 was taken in order to calculate the β-galactosidase activity with the equation 1000 x OD420/(t x V x OD600) where t is the elapsed time (minutes) and V is the amount of lysate used.

The specificity of the HIP1-HD interaction can be observed using the chromogenic filter assay. Only yeast cells harboring HIP1 and HD activate both the HIS and lacZ reporter genes in the Y190 yeast host. The cells that contain the HIP1 with HD constructs with 80 or 128 CAG repeats turn blue approximately 45 minutes after the cells with the smaller sized repeats (16 or 44).

No difference in the β-galactosidase activity was observed between the 16 and 44 repeats or between the 80 and 128 repeats. However, a significant difference ($p<0.05$) in activity is seen between the smaller repeats (16 and 44) and the larger repeats (80 and 128). (Figure 1)

EXAMPLE 3

DNA SEQUENCING, cDNA ISOLATION AND 5' RACE

Oligonucleotide primers were synthesized on an ABI PCR-mate oligo-synthesizer. DNA sequencing was performed using an ABI 373 fluorescent automated DNA sequencer. The HIP cDNAs were confirmed to be in-frame with the GAL4-AD by sequencing across the AD-HIP1 cloning junction using an AD oligonucleotide (5'GAA GAT ACC CCA CCA AAC3').

Subsequently, primer walking was used to determine the remaining sequences. A human frontal cortex >4.0 kb cDNA library (a gift from S. Montal) was screened to isolate the full length HIP1 gene. Fifty nanograms of a 558 base pair Eco RI fragment from the original HIP1 cDNA was radioactively labeled with [α^{32}P]-dCTP using nick-translation and the probe allowed to hybridized to filters containing > 105 pfu/ml of the cDNA library overnight at 65 C in Church buffer (see Northern blot protocol). The filters were washed at 65 C for 10 minutes with 1 X SSPE, 15 minutes at 65 C with 1 X SSPE and 0.1% SDS, then for thirty minutes and fifteen minutes with 1 X SSPE and 0.1% SDS. The filters were exposed to X-ray film (Kodak, XAR5) overnight at -70 C. Primary

positives were isolated and replated and subsequent secondary positives were hybridized and washed as for the primary screen. The resulting positive phage were converted into plasmid DNA by conventional methods (Stratagene) and the cDNA isolated and sequenced.

In order to obtain the most 5' sequence of the HIP1 gene, a Rapid Amplification of cDNA Ends (RACE) protocol was performed according to the manufacturers recommendations (BRL). First strand cDNA was synthesized using the oligo HIP1-242R (5' GCT TGA CAG TGT AGT CAT AAA GGT GGC TGC AGT CC 3'). After dCTP tailing the cDNA with terminal deoxy transferase, two rounds of 35 cycles (94 C 1 minute; 53 C 1 minute; 72 C 2 minutes) of PCR using HIP1-R2 (5' GGA CAT GTC CAG GGA GTT GAA TAC 3') and an anchor primer (5' (CUA)4 GGC CAC GCG TCG ACT AGT ACG GGI IGG GII GGG IIG3') (BRL) were performed. The subsequent 650 base pair PCR product was cloned using the TA cloning system (Invitrogen) and sequenced using T3 and T7 primers. Sequences ID Nos. 1 and 5 show the sequence of the HIP1 cDNAs obtained.

EXAMPLE 4

DNA AND AMINO ACID ANALYSES

Overlapping DNA sequence was assembled using the program MacVector and sent via email or Netscape to the BLAST server at NIH (http://www.ncbi.nlm.nih.gov) to search for sequence similarities with known DNA (blastn) or protein (tblastn) sequences. Amino acid alignments were performed with the program Clustalw.

EXAMPLE 5

FISH DETECTION SYSTEM AND IMAGE ANALYSIS

The HIP1 cDNA isolated from the two-hybrid screen was mapped by fluorescent in situ hybridization (FISH) to normal human lymphocyte chromosomes counterstained with propidium iodide and DAPI. Biotinylated probe was detected with avidin-fluorescein isothiocyanate (FITC). Images of metaphase preparations were captured by a thermoelectrically cooled charge coupled camera (Photometrics). Separate images of DAPI banded chromosomes and FITC targeted chromosomes were obtained. Hybridization

signals were acquired and merged using image analysis software and pseudo colored blue (DAPI) and yellow (FITC) as described and overlaid electronically. This study showed that HIP1 maps to a single genomic locus at 7q11.2.

EXAMPLE 6

NORTHERN BLOT ANALYSIS

RNA was isolated using the single step method of homogenization in guanidinium isothiocyante and fractionated on a 1.0% agarose gel containing 0.6 M formaldehyde. The RNA was transferred to a hybond N -membrane (Amersham) and crosslinked with ultraviolet radiation.

Hybridization of the Northern blot with b-actin as an internal control probe provided confirmation that the RNA was intact and had transferred. The 1.2 kb HIP1 cDNA was labeled using nick translation and incorporation of α^{32}P-dCTP. Hybridization of the original 1.2 kb HIP1 cDNA was carried out in Church buffer (0.5 M sodium phosphate buffer, pH 7.2, 2.7% sodium dodecyl sulphate, 1 mM EDTA) at 55 C over-night. Following hybridization, Northern blots were washed once for 10 minutes in 2.0 X SSPE, 0.1% SDS at room temperature and twice for 10 minutes in 0.15 X SSPE, 0.1% SDS. Autoradiography was carried our from one to three days using Hyperfilm (Amersham) film at -70 C.

Analysis of the levels of RNA levels of HIP1 by Northern blot data revealed that the 10 kilo base HIP1 message is present in all tissue assessed. However, the levels of RNA are not uniform, with brain having highest levels of expression and peripheral tissues having less message. No apparent differences in RNA expression was noted between control samples and HD affected individuals.

EXAMPLE 7

TISSUE LOCALIZATION OF HIP1

Tissue localization of HIP1 was studied using a variety of techniques as described below. Subcellular distribution of HIP-1 protein in adult human and mouse brain Biochemical fractionation studies revealed the HIP1 protein was found to be a

membrane-associated protein. No immunoreactivity was seen by Western blotting in cytosolic fractions, using the anti-HIP1-pep1 polyclonal antibody. HIP1 immunoreactivity was observed in all membrane fractions including nuclei (P1), mitochondria and synaptosomes (P2), microsomes and plasma membranes (P3). The P3 fraction contained the most HIP1 compared to other membrane fractions. HIP1 could be removed from membranes by high salt (0.5M NaCl) buffers indicating it is not an integral membrane protein, however, since low salt (0.1- 0.25M NaCl) was only able to partially remove HIP1 from membranes, its membrane association is relatively strong. The extraction of P3 membranes with the non-ionic detergent, Triton X-100 revealed HIP1 to be a Triton X-100 insoluble protein. This characteristic is shared by many cytoskeletal and cytoskeletal-associated membrane proteins including actin, which was used as a control in this study. The biochemical characteristics of HIP1 described were found to be identical in mouse and human brain and was the same for both forms of the protein (both bands of the HIP1 doublet). HIP1 co-localized with huntingtin in the P2 and P3 membrane fractions, including the high-salt membrane extractions, as well as in the Triton X-100 insoluble residue. The subcellular distribution of HIP1 was unaffected by the expression of polyglutamine-expanded huntingtin in transgenic mice and HD patient brain samples.

The localization of HIP1 protein was further investigated by immunohistochemistry in normal adult mouse brain tissue. Immunoreactivity was seen in a patchy, reticular pattern in the cytoplasm, appeared excluded from the nucleus and stained most intensely in a discontinuous pattern at the membrane. These results are consistent with the association of HIP1 with the cytoskeletal matrix and further indicate an enrichment of HIP1 at plasma membranes. Immunoreactivity occurred in all regions of the brain, including cortex, striatum, cerebellum and brainstem, but appeared most strongly in neurons and especially in cortical neurons. As described previously, huntingtin immunoreactivity was seen exclusively and uniformly in the cytosol.

The in situ hybridization studies showed HIP1 mRNA to be ubiquitously and generally expressed throughout the brain. This data is consistent with the immunohistochemical results and was identical to the distribution pattern of huntingtin mRNA in transgenic mouse brains expressing full-length human huntingtin.

Protein Preparation And Western Blotting For Expression Studies

Frozen human tissues were homogenized using a Polytron in a buffer containing 0.25M sucrose, 20mM Tris-HCl (pH 7.5), 10mM EGTA, 2mM EDTA supplemented with 10ug/ml of leupeptin, soybean trypsin inhibitor and 1mM PMSF, then centrifuged at 4,000rpm for 10' at 4 C to remove cellular debris. 100-150ug/lane of protein was separated on 8% SDS-PAGE mini-gels and then transferred to PVDF membranes. Huntingtin and HIP1 were electroblotted overnight in Towbin's transfer buffer (25 mM Tris-HCl, 0.192M glycine, pH8.3, 10% methanol) at 30V onto PVDF membranes (Immobilon-P, Millipore) as described (Towbin et al, *Proc. Nat'l Acad. Sci.(USA)* 76: 4350-4354 (1979)). Membranes were blocked for 1 hour at room temperature in 5% skim milk/ TBS (10mM Tris-HCl, 0.15M NaCl, pH7.5). Antibodies against huntingtin (pAb BKP1, 1:500), actin (mAb A-4700, Sigma, 1:500) or HIP1 (pAb HIP-pep1, 1:200) were added to blocking solution for 1 hour at room temperature. After 3 x 10 minutes washes in TBS-T (0.05% Tween-20/TBS), secondary Ab (horseradish peroxidase conjugated IgG, Biorad) was applied in blocking solution for 1 hour at room temperature. Membranes were washed and then incubated in chemiluminescent ECL solution and visualized using Hyperfilm-ECL film (Amersham).

Generation of Antibodies

The generation of huntingtin specific antibodies GHM1 and BKP1 is described elsewhere (Kalchman, et al., *J. Biol. Chem.* 271: 19385-19394 (1996)). The HIP1 peptide (VLEKDDLMDMDASQQN, a.a. 76-91 of Seq. ID No. 2) was synthesized with Cys on the N-terminus for the coupling, and coupled to Keyhole limpet hemocyanin (KLH) (Pierce) with succinimidyl 4-(N-maleimidomethyl) cyclohexame-1-carboxylate (Pierce). Female New Zealand White rabbits were injected with HIP1 peptide-KLH and Freund's adjuvant. Antibodies against the HIP1 peptide were purified from rabbit sera using affinity column with low pH elution. Affinity column was made by incubation of HIP1 peptide with activated thio-Sepharose (Pharmacia).

Western blotting of various peripheral and brain tissues were consistent with the RNA data. The HIP1 protein levels observed was not ubiquitous. The protein

expression is limited to brain tissue, with highest amounts seen in the cortex and lower levels seen in the cerebellum and caudate and putamen.

More regio-specific analysis of HIP1 expression in the brain revealed no differential expression pattern in affected individuals when compared to normal controls, with highest levels of expression seen in both controls and HD patients in the cortical regions.

EXAMPLE 8

CO-IMMUNOPRECIPITATION OF HIP1 WITH HUNTINGTIN

Confirmation of the HD-HIP1 interaction was performed using coimmunoprepitation as follows. Control human brain (frontal cortex) lysate was prepared in the same manner as for subcellular localization study. Prior to immunoprecipitation, tissue lysate was centrifuged at 5000 rpm for 2 minutes at 4 C, then the supernatant was pre-cleared by the incubated with excess amount of Protein A-Sepharose for 30 minutes at 4 C, and centrifuged at the same condition. Fifty microlitres of supernatant (500 mg protein) was incubated with or without antibodies (10 ug of anti-huntingtin GHM1 (Kalchman, et al. 1996) or anti-synaptobrevin antibody) in the total 500 ul of incubation buffer (20mM Tris-Cl (pH7.5), 40mM NaCl, 1mM $MgCl_2$) for 1 hour at 4 C. Twenty microlitres of Protein A-Sepharose (1:1 suspension, for GHM1 and no antibody control) or Protein G-Sepharose (for anti-synaptobrevin antibody; Pharmacia) was added and incubated for 1 hour at 4 C. The beads were washed with washing buffer (incubation buffer containing 0.5 % Triton X-100) three times. The samples on the beads were separated using SDS-PAGE (7.5% acrylamide) and transferred to PVDF membrane (Immobilon-P, Millipore). The membrane was cut at about 150 kDa after transfer for Western blotting (as described above). The upper piece was probed with anti-huntingtin BKP1 (1/1000) and lower piece with anti-HIP1 antibody (1/300).

The results showed that when an anti-HIP1 polyclonal antibody was immunoreacted against a blot containing the GHM1 immunoprecipitates from the brain lysate a doublet was observed at approximately 100 kDa was. When GHM1 was immunoreacted against the same immunoprecipitate the 350 kDa HD protein was also seen The

specificity of the HD-HIP1 interaction is seen as no immunoreactive bands seen are as a result of the proteins adsorbing to the Protein-A-Sepharose (Lysate + No Antibody) or when a random, non related antibody (Lysate + anti-Synaptobrevin) is used as the immunoprecipitating antibody.

EXAMPLE 9

Subcellular fractionation of brain tissue

Cortical tissue (20-100 mg/ml) was homogenized, on ice, in a 2 ml pyrex-teflon IKA-RW15 homogenizer (Tekmar Company) in a buffer containing 0.303M sucrose, 20mM Tris-HCl pH 6.9, 1mM $MgCl_2$, 0.5mM EDTA, 1mM PMSF, 1mM leupeptin, soybean trypsin inhibitor and 1mM benzamidine (Wood et al., *Human Molec. Genet.* 5: 481-487 (1996)).

Crude membrane vesicles were isolated by two cycles of a three-step differential centrifugation protocol in a Beckman TLA 120.2 rotor at 4 C based on the methods of Wood et al (1996). The first step precipitated cellular debris and nuclei from tissue homogenates for 5 minutes at 1300 x g (P1). The 1300 x g supernatant was subsequently centrifuged for 20 minutes at 14 000 x g to isolate synaptosomes and mitochondria (P2). Finally, microsomal and plasma membrane vesicles were collected by a 35 minute centrifugation at 142 000 x g (P3). The remaining supernatant was defined as the cytosolic fraction.

High salt extraction of membranes

Aliquots of P3 membranes were twice suspended at 2mg/ ml in 0.5M NaCl, 10mM Tris-HCl, 2mM $MgCl_2$, pH7.2, containing protease inhibitors (see above). The same buffer without NaCl was used as a control. The membrane suspensions were incubated on ice for 30 minutes and then centrifuged at 142 000 x g for 30 minutes.

Extraction of cytoskeletal and cytoskeletal-associated proteins.

To extract cytoskeletal proteins, crude membrane vesicles from the P3 fraction membrane were suspended in a volume of Triton X-100 extraction buffer to give a

protein: detergent ratio of 5:1. The composition of the Triton X-100 extraction buffer was based on the methods of Arai et al., *J. Neuroscience* 38: 348-357 (1994) and contained 2% Triton X-100, 10mM Tris-HCl, 2mM $MgCl_2$, 1mM leupeptin, soybean trypsin inhibitor, PMSF and benzamidine. Membrane pellets were suspended by hand with a round-bottom teflon pestle, and placed on ice for 40 minutes. Insoluble cytoskeletal matrices were precipitated for 35 minutes at 142 000 x g in a Beckman TLA 120.2 rotor. The supernatant was defined as non-cytoskeletal-associated membrane or membrane--associated protein and was removed. The remaining pellet was extracted with Triton X-100 a second time using the same conditions. We defined the final pellet as cytoskeletal and cytoskeletal-associated protein.

Solubilization of protein and analysis by SDS-PAGE and Western Blotting

Membrane and cytoskeletal protein was solubilized in a minimum volume of 1% SDS, 3M urea, 0.1mM dithiothreitol in TBS buffer and sonicated. Protein concentration was determined using the BioRad DC Protein assay and samples were diluted at least 1 X with 5 X sample buffer (250mM Tris-HCl pH 6.8, 10% SDS, 25% glycerol, 0.02% bromophenol blue and 7% 2-mercaptoethanol) and were loaded on 7.5% SDS-PAGE gels (Bio-Rad Mini-PROTEIN II Cell system) without boiling. Western blotting was performed as described above.

Immunohistochemistry

Brain tissue was obtained from a normal C57BL/6 adult (6 months old) male mouse sacrificed with chloroform then perfusion-fixed with 4% v/v paraformaldehyde/0.01 M phosphate buffer (4% PFA). The brain tissues were removed, immersion fixed in 4% PFA for 1 day, washed in 0.01M phosphate buffered saline, pH 7.2 (PBS) for 2 days, and then equilibrated in 25% w/v sucrose PBS for 1 week. The samples were then snap-frozen in Tissue Tek molds by isopentane cooled in liquid nitrogen. After warming to -20 C, frozen blocks derived from frontal cortex, caudate/putamen, cerebellum and brainstem were cut into 14 mm sections for immunohistochemistry. Following washing in PBS, the tissue sections were blocked using 2.5% v/v normal goat serum for 1 hour at room

temperature. Primary antibodies diluted with PBS were applied to sections overnight at 4 C. Optimal dilutions for the polyclonal antibodies BKP1 and HIP1 were 1:50. Using washes of 3 x 5 minutes in PBS at room temperature, sections were sequentially incubated with biotinylated secondary antibody and then an avidin-biotin complex reagent (Vecta Stain ABC Kit, Vector) for 60 minutes each at room temperature. Color was developed using 3-3'-diaminobenzidine tetrahydrocholoride and ammonium nickel sulfate.

For controls, sections were treated as described above except that HIP1 antibody aliquots were preabsorbed with an excess of HIP1 peptide as well as a peptide unrelated to HIP1 prior to incubation with the tissue sections.

In situ hybridization

In situ hybridization was performed as previously described with some modification (Suzuki et al, *BBRC* 219: 708-713 (1996)). The RNA probes were prepared using the plasmid gt149 (Lin, B., et al., *Human Molec. Genet.* 2: 1541-1545 (1994)) or a 558 subclone of HIP1. The anti-sense and sense single-stranded RNA probes were synthesized using T3 and T7 RNA polymerases and the In Vitro Transcription Kit (Clontech) with the addition of $[\alpha^{35}S]$-CTP (Amersham) to the reaction mixture. Sense RNA probes were used as negative controls. For HIP1 studies normal C57BL/6 mice were used. Huntingtin probes were tested on two different transgenic mouse strains expressing full-length huntingtin, cDNA HD10366(44CAG) C57BL/6 mice and YAC HD10366(18CAG) FVB/N mice. Frozen brain sections (10um thick) were placed onto silane-coated slides under RNase-free conditions. The hybridization solution contained 40% w/v formamide, 0.02M Tris-HCl (pH 8.0), 0.005M EDTA, 0.3 M NaCl, 0.01M sodium phosphate (pH 7.0), 1x Denhardt's solution, 10% w/v dextran sulfate (pH 7.0), 0.2% w/v sarcosyl, yeast tRNA (500mg/ml) and salmon sperm DNA (200mg/ml). The radiolabelled RNA probe was added to the hybridization solution to give 1 x 106 cpm/200 ul/ section. Sections were covered with hybridization solution and incubated on formamide paper at 65 C for 18 hours. After hybridization, the slides were washed for 30 minutes sequentially with 2x SSC, 1x SSC and high stringency wash solution (50% formamide, 2x SSC and 0.1M dithiothreitol) at 65 C, followed by treatment with Rnase A

(1mg/ml) at 37 C for 30 minutes, then washed again and air-dried. The slides were first exposed on autoradiographic film (b-max, Amersham, UK) for 48 hours and developed for 4 minutes in Kodak D-19 followed by a 5 minute fixation in Fuji-fix. For longer exposures, the slides were dipped in autoradiographic emulsion (50% w/v in distilled water, NR-2, Konica, Japan), air-dried and exposed for 20 days at 4 C then developed as described. Sections were counterstained with methyl green or Giemsa solutions.

SEQUENCE LISTING

(1) GENERAL INFORMATION:
(i) APPLICANT: Kalchman, Michael
Goldberg, Paul
Hayden. Michael R.
(ii) TITLE OF INVENTION: Protein Which Interacts with the Huntington's Disease Gene Product, cDNA Coding Therefor, and Antibodies Thereto
(iii) NUMBER OF SEQUENCES: 8
(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Oppedahl & Larson
(B) STREET: 1992 Commerce Street Suite 309
(C) CITY: Yorktown
(D) STATE: NY
(E) COUNTRY: USA
(F) ZIP: 10598
(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Kb storage
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: MS DOS 5.0
(D) SOFTWARE: WordPerfect
(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:
(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Larson, Marina T.
(B) REGISTRATION NUMBER: 32038
(C) REFERENCE/DOCKET NUMBER: UBC.P-013
(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (914) 245-3252
(B) TELEFAX: (914) 962-4330

(2) INFORMATION FOR SEQ ID NO:1:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1164
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii)MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: no
(iv) ANTI-SENSE: no
(vi) ORIGINAL SOURCE:
(A) ORGANISM: human
(ix) FEATURE: cDNA for Huntingtin-interacting protein
(xi)SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
ACAGCTGACA CCCTGCAAGG CCACCGGGAC CGCTTCATGG AGCAGTTTAC     50
AAAGTTGAAA GATCTGTTCT ACCGCTCCAG CAACCTGCAG TACTTCAAGC    100
GGGTCATTCA GATCCCCCAG CTGCCTGAGA ACCCACCCAA CTTCCTGCGA    150
GCCTCAGCCC TGTCAGAACA TATCAGCCCT GTGGTGGTGA TCCCTGCAGA    200
GGCCTCATCC CCCGACAGCG AGCCAGTCCT AGAGAAGGAT GACCTCATGG    250
ACATGGATGC CTCTCAGCAG AATTTATTTG ACAACAAGTT TGATGACNTC    300
TTTGGCAGTT CATCCAGCAG TGATCCCTTC AATTTCAACA GTCAAAATGG    350
TGTGAACAAG GATGAGAAGG ACCACTTAAT TGAGCGACTA TACAGAGAGA    400
TCAGTGGATT GAAGGCACAG CTAGAAAACA TGAAGACTGA GAGCCAGCGG    450
GTTGTGCTGC AGCTGAAGGG CCACGTCAGC GAGCTGGAAG CAGATCTGGC    500
CGAGCAGCAG CACCTGCGGC AGCAGGCGGC CGACGACTGT GAATTCCTGC    550
GGGCAGAACT GGACGAGCTC AGGNGGCAGC GGGAGGACAC CGAGAAGGCT    600
CAGCGGAGCC TGTCTGAGAT AGAAAGGAAA GCTCAAGCCA ATGAACAGCG    650
ATATAGCAAG CTAAAGGAGA AGTACAGCGA GCTGGTTCAG AACCACGCTG    700
ACCTGCTGCG GAAGAATGCA GAGGTGACCA AACAGGTGTC CATGGCCAGA    750
CAAGCCCAGG TAGATTTGGA ACGAGAGAAA AAAGAGCTGG AGGATTCGTT    800
GGAGCGCATC AGTGACCAGG GCCAGCGGAA GACTCAAGAA CAGCTGGAAG    850
TTCTAGAGAG CTTGAAGCAG GAACTTGGCA CAAGCCAACG GGAGCTTCAG    900
GTTCTGCAAG GCAGCCTGGA AACTTCTGCC CAGTCAGAAG CAAACTGGGC    950
AGCCGAGTTC GCCGAGCTAG AGAAGGAGCG GGACAGCCTG GTGAGTGGCG   1000
CAGCTCATAG GGAGGAGGAA TTATCTGCTC TTCGGAAAGA ACTGCAGGAC   1050
ACTCAGCTCA AACTGGCCAG CACAGAGGAA TCTATGTGCC AGCTTGCCAA   1100
AGACCAACGA AAAATGCTTC TGGTGGGGTC CAGGAAGGCT GCGGAGCAGG   1150
TGATACAAGA CGCG                                          1164
```

(2) INFORMATION FOR SEQ ID NO:2:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 386
(B) TYPE: protein
(D) TOPOLOGY: linear
(ii)MOLECULE TYPE: protein
(iii) HYPOTHETICAL: no
(vi) ORIGINAL SOURCE:
(A) ORGANISM: human
(ix) FEATURE: Huntingtin-interacting protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Thr Ala Asp Thr Leu Gln Gly His Arg Asp Arg Phe Met Glu Gln
 1               5                  10                  15

Phe Thr Lys Leu Lys Asp Leu Phe Tyr Arg Ser Ser Asn Leu Gln
                20                  25                  30

Tyr Phe Lys Arg Val Ile Gln Ile Pro Gln Leu Pro Glu Asn Pro
                35                  40                  45

Pro Asn Phe Leu Arg Ala Ser Ala Leu Ser Glu His Ile Ser Pro
                50                  55                  60

Val Val Val Ile Pro Ala Glu Ala Ser Ser Pro Asp Ser Glu Pro
                65                  70                  75
```

Val Leu Glu Lys Asp Asp Leu Met Asp Met Asp Ala Ser Gln Gln
80 85 90

Asn Leu Phe Asp Asn Lys Phe Asp Asp Phe Gly Ser Ser Ser Ser
95 100 105

Ser Asp Pro Phe Asn Phe Asn Ser Gln Asn Gly Val Asn Lys Asp
110 115 120

Glu Lys Asp His Leu Ile Glu Arg Leu Tyr Arg Glu Ile Ser Gly
125 130 135

Leu Lys Ala Gln Leu Glu Asn Met Lys Thr Glu Ser Gln Arg Val
140 145 150

Val Leu Gln Leu Lys Gly His Val Ser Glu Leu Glu Ala Asp Leu
155 160 165

Ala Glu Gln Gln His Leu Arg Gln Gln Ala Ala Asp Asp Cys Glu
170 175 180

Phe Leu Arg Ala Glu Leu Asp Glu Leu Arg Gln Arg Glu Asp Thr
185 190 195

Glu Lys Ala Gln Arg Ser Leu Ser Glu Ile Glu Arg Lys Ala Gln
200 205 210

Ala Asn Glu Gln Arg Tyr Ser Lys Leu Lys Glu Lys Tyr Ser Glu
215 220 225

Leu Val Gln Asn His Ala Asp Leu Leu Arg Lys Asn Ala Glu Val
230 235 240

Thr Lys Gln Val Ser Met Ala Arg Gln Ala Gln Val Asp Leu Glu
245 250 255

Arg Glu Lys Lys Glu Leu Glu Asp Ser Leu Glu Arg Ile Ser Asp
260 265 270

Gln Gly Gln Arg Lys Thr Gln Glu Gln Leu Glu Val Leu Glu Ser
275 280 285

Leu Lys Gln Glu Leu Gly Thr Ser Gln Arg Glu Leu Gln Val Leu
290 295 300

Gln Gly Ser Leu Glu Thr Ser Ala Gln Ser Glu Ala Asn Trp Ala
305 310 315

Ala Glu Phe Ala Glu Leu Glu Lys Glu Arg Asp Ser Leu Val Ser
320 325 330

Gly Ala Ala His Arg Glu Glu Glu Leu Ser Ala Leu Arg Lys Glu
335 340 345

```
Leu Gln Asp Thr Gln Leu Lys Leu Ala Ser Thr Glu Glu Ser Met
                350                 355                 360

Cys Gln Leu Ala Lys Asp Gln Arg Lys Met Leu Leu Val Gly Ser
                365                 370                 375

Arg Lys Ala Ala Glu Gln Val Ile Gln Asp Ala
                380                 385 386
```

(2) INFORMATION FOR SEQ ID NO:3:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH:
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii)MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: no
(iv) ANTI-SENSE: no
(vi) ORIGINAL SOURCE:
(A) ORGANISM: human
(ix) FEATURE: cDNA for Huntingtin-interacting protein
(xi)SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
ACCGATACCG AAGCGGGCTG TGTGCCCCTT CTCCACCCAG AGGAAATCAA    50
ACCCCAAAGC CATTATAACC ATGGATATGG TGAACCTCTT GGACGGAAAA   100
CTCATATTGA TGATTACAGC ACATGGGACA TAGTCAAGGC TACACAATAT   150
GGAATATATG AACGCTGTCG AGAATTGGTG GAAGCAGGTT ATGATGTACG   200
GCAACCGGAC AAAGAAAATG TTACCCTCCT CCATTGGGCT GCCATCAATA   250
ACAGAATAGA TTTAGTCAAA TACTATATTT CGAAAGGTGC TATTGTGGAT   300
CAACTTGGAG GGGACCTGAA TTCAACTCCA TTGCACTGGG ACACAAGACA   350
AGGCCATCTA TCCATGGTTG TGCAACTAAT GAAATATGGT GCAGATCCTT   400
CATTAATTGA TGGAGAAGGA TGTAGCTGTA TTCATCTGGC TGCTCAGTTC   450
GGACATACCT CAATTGTTGC TTATCTCATA GCAAAAGGAC AGGATGTG      498
```

(2) INFORMATION FOR SEQ ID NO:4:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 154
(B) TYPE: protein
(D) TOPOLOGY: linear
(ii)MOLECULE TYPE: protein
(iii) HYPOTHETICAL: no
(vi) ORIGINAL SOURCE:
(A) ORGANISM: human
(ix) FEATURE: Huntingtin-interacting protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Thr Asp Thr Glu Ala Gly Cys Val Pro Leu Leu His Pro Glu Glu
 1               5                  10                  15

Ile Lys Pro Gln Ser His Tyr Asn His Gly Tyr Gly Glu Pro Leu
                20                  25                  30

Gly Arg Lys Thr His Ile Asp Asp Tyr Ser Thr Trp Asp Ile Val
                35                  40                  45

Lys Ala Thr Gln Tyr Gly Ile Tyr Glu Arg Cys Arg Glu Leu Val
                50                  55                  60

Glu Ala Gly Tyr Asp Val Arg Gln Pro Asp Lys Glu Asn Val Thr
                65                  70                  75

Leu Leu His Trp Ala Ala Ile Asn Asn Arg Ile Asp Leu Val Lys
                80                  85                  90

Tyr Tyr Ile Ser Lys Gly Ala Ile Val Asp Gln Leu Gly Gly Asp
                95                 100                 105

Leu Asn Ser Thr Pro Leu His Trp Asp Thr Arg Gln Gly His Leu
               110                 115                 120

Ser Met Val Val Gln Leu Met Lys Tyr Gly Ala Asp Pro Ser Leu
               125                 130                 135

Ile Asp Gly Glu Gly Cys Ser Cys Ile His Leu Ala Ala Gln Phe
               140                 145                 150

Gly His Thr Ser
           154
```

(2) INFORMATION FOR SEQ ID NO:5:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4846
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii)MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: no
(iv) ANTI-SENSE: no
(vi) ORIGINAL SOURCE:
(A) ORGANISM: human
(ix) FEATURE: cDNA for Huntingtin-interacting protein
(xi)SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
CAGTGTACGG TTGATCATAT AACGCCGCGG GCGGGGATTG GTTTATATAT    50
CGCAAATTGA TNTAGGGGGG GGGGGATGGN CAGAGATTTC GCTTCATTAG   100
GCCATTATAA GCAGGAAGGG TTTCAAGGAA AAAAACCCAG AAAGTGCATA   150
TTGCACCCAC CATGAGAAAG GGGCAACAGA CCTTNTGTTN TGTTNTCAAC   200
```

```
CGCCTGCTTC TGTTTTAGCA ACGCAGTGTT TTGGTGGAAG TTGTGCCATG   250
TGTTCCACAA ANTCTTCCGA GATGGACACC CGAACGTCCT GAAGGACTTT   300
GTGAGATACA GAAATGAATT GAGTGACATG AGCAGGATGT GGGGCCACCT   350
GAGCGAGGGG TATGGCCAGC TGTGCAGCAT CTACCTGAAA CTGCTAAGAA   400
CCAAGATGGA GTACCACACC AAAAATCCCA GGTTCCCAGG CAACCTGCAG   450
ATGAGTGACC GCCAGCTGGA CGAGGCTGGA GAAAGTGACG TGAACAACTT   500
TTTCCAGTTA ACAGTGGAGA TGTTTGACTA CCTGGAGTGT GAACTCAACC   550
TCTTCCAAAC AGTATTCAAC TCCCTGGACA TGTCCCGCTC TGTGTCCGTG   600
ACGGCAGCAG GGCAGTGCCG CCTCGCCCCG CTGATCCAGG TCATCTTGGA   650
CTGCAGCCAC CTTTATGACT ACACTGTCAA GCTTCTCTTC AAACTCCACT   700
CCTGCCTCCC AGCTGACACC CTGCAAGGCC ACCGGGACCG CTTCATGGAG   750
CAGTTTACAA AGTTGAAAGA TCTGTTCTAC CGCTCCAGCA ACCTGCAGTA   800
CTTCAAGCGG CTCATTCAGA TCCCCCAGCT GCCTGAGAAC CCACCCAACT   850
TCCTGCGAGC CTCAGCCCTG TCAGAACATA TCAGCCCTGT GGTGGTGATC   900
CCTGCAGAGG CCTCATCCCC CGACAGCGAG CCAGTCCTAG AGAAGGATGA   950
CCTCATGGAC ATGGATGCCT CTCAGCAGAA TTTATTTGAC AACAAGTTTG  1000
ATGACATCTT TGGCAGTTCA TTCAGCAGTG ATCCCTTCAA TTTCAACAGT  1050
CAAAATGGTG TGAACAAGGA TGAGAAGGAC CACTTAATTG AGCGACTATA  1100
CAGAGAGATC AGTGGATTGA AGGCACAGCT AGAAAACATG AAGACTGAGA  1150
GCCAGCGGGT TGTGCTGCAG CTGAAGGGCC ACGTCAGCGA GCTGGAAGCA  1200
GATCTGGCCG AGCAGCAGCA CCTGCGGCAG CAGGCGGCCG ACGACTGTGA  1250
ATTCCTGCGG GCAGAACTGG ACGAGCTCAG GAGGCAGCGG GAGGACACCG  1300
AGAAGGCTCA GCGGAGCCTG TCTGAGATAG AAAGGAAAGC TCAAGCCAAT  1350
GAACAGCGAT ATAGCAAGCT AAAGGAGAAG TACAGCGAGC TGGTTCAGAA  1400
CCACGCTGAC CTGCTGCGGA AGAATGCAGA GGTGACCAAA CAGGTGTCCA  1450
TGGCCAGACA AGCCCAGGTA GATTTGGAAC GAGAGAAAAA AGAGCTGGAG  1500
GATTCGTTGG AGCGCATCAG TGACCAGGGC CAGCGGAAGA CTCAAGAACA  1550
GCTGGAAGTT CTAGAGAGCT TGAAGCAGGA ACTTGGCACA AGCCAACGGG  1600
AGCTTCAGGT TCTGCAAGGC AGCCTGGAAA CTTCTGCCCA GTCAGAAGCA  1650
AACTGGGCAG CCGAGTTCGC CGAGCTAGAG AAGGAGCGGG ACAGCCTGGT  1700
GAGTGGCGCA GCTCATAGGG AGGAGGAATT ATCTGCTCTT CGGAAAGAAC  1750
TGCAGGACAC TCAGCTCAAA CTGGCCAGCA CAGAGGAATC TATGTGCCAG  1800
CTTGCCAAAG ACCAACGAAA AATGCTTCTG GTGGGGTCCA GGAAGGCTGC  1850
GGAGCAGGTG ATACAAGACG CCCTGAACCA GCTTGAAGAA CCTCCTCTCA  1900
TCAGCTGCGC TGGGTCTGCA GATCACCTCC TCTCCACGGT CACATCCATT  1950
TCCAGCTGCA TCGAGCAACT GGAGAAAAGC TGGAGCCAGT ATCTGGCCTG  2000
CCCAGAAGAC ATCAGTGGAC TTCTCCATTC CATAACCCTG CTGGCCCACT  2050
TGACCAGCGA CGCCATTGCT CATGGTGCCA CCACCTGCCT CAGAGCCCCA  2100
CCTGAGCCTG CCGACTCACT GACCGAGGCC TGTAAGCAGT ATGGCAGGGA  2150
AACCCTCGCC TACCTGGCCT CCCTGGAGGA AGAGGGAAGC CTTGAGAATG  2200
CCGACAGCAC AGCCATGAGG AACTGCCTGA GCAAGATCAA GGCCATCGGC  2250
GAGGAGCTCC TGCCCAGGGG ACTGGACATC AAGCAGGAGG AGCTGGGGGA  2300
CCTGGTGGAC AAGGAGATGG CGGCCACTTC AGCTGCTATT GAAACTTGCA  2350
CGGCCAGAAT AGAGGAGATG CTCAGCAAAT CCCGAGCAGG AGACACAGGA  2400
GTCAAATTGG AGGTGAATGA AAGGATCCTT CGTTGCTGTA CCAGCCTCAT  2450
GCAAGCTATT CAGGTGCTCA TCGTGGCCTC TAAGGACCTC CAGAGAGAGA  2500
TTGTGGAGAG CGGCAGGGGT ACAGCATCCC CTAAAGAGTT TTATGCCAAG  2550
AACTCTCGAT GGACAGAAGG ACTTATCTCA GCCTCCAAGG CTGTGGGCTG  2600
GGGAGCCACT GTCATGGTGG ATGCAGCTGA TCTGGTGGTA CAAGGCAGAG  2650
GGAAATTTGA GGAGCTAATG GTGTGTTCTC ATGAAATTGC TGCTAGCACA  2700
GCCCAGCTTG TGGCTGCATC CAAGGTGAAA GCTGATAAGG ACAGCCCCAA  2750
CCTAGCCCAG CTGCAGCAGG CCTCTCGGGG AGTGAACCAG GCCACTGCCG  2800
GCGTTGTGGC CTCAACCATT TCCGGCAAAT CACAGATCGA AGAGACAGAC  2850
AACATGGACT TCTCAAGCAT GACGCTGACA CAGATCAAAC GCCAAGAGAT  2900
```

GGATTCTCAG GTTAGGGTGC TAGAGCTAGA AAATGAATTG CAGAAGGAGC 2950
GTCAAAAACT GGGAGAGCTT CGGAAAAAGC ACTACGAGCT TGCTGGTGTT 3000
GCTGAGGGCT GGGAAGAAGG AACAGAGGCA TCTCCACCTA CACTGCAAGA 3050
AGTGGTAACC GAAAAAGAAT AGAGCCAAAC CAACACCCCA TATGTCAGTG 3100
TAAATCCTTG TTACCTATCT CGTGTGTGTT ATTTCCCCAG CCACAGGCCA 3150
AATCCTTGGA GTCCCAGGGG CAGCCACACC ACTGCCATTA CCCAGTGCCG 3200
AGGACATGCA TGACACTTCC CAAAGATCCC TCCATAGCGA CACCCTTTCT 3250
GTTTGGACCC ATGGTCATCT CTGTTCTTTT CCCGCCTCCC TAGTTAGCAT 3300
CCAGGCTGGC CAGTGCTGCC CATGAGCAAG CCTAGGTACG AAGAGGGGTG 3350
GTGGGGGGCA GGGCCACTCA ACAGAGAGGA CCAACATCCA GTCCTGCTGA 3400
CTATTTGACC CCCACAACAA TGGGTATCCT TAATAGAGGA GCTGCTTGTT 3450
GTTTGTTGAC AGCTTGGAAA GGGAAGATCT TATGCCTTTT CTTTTCTGTT 3500
TTCTTCTCAG TCTTTTCAGT TTCATCATTT GCACAAACTT GTGAGCATCA 3550
GAGGGCTGAT GGATTCCAAA CCAGGACACT ACCCTGAGAT CTGCACAGTC 3600
AGAAGGACGG CAGGAGTGTC CTGGCTGTGA ATGCCAAAGC CATTCTCCCC 3650
CTCTTTGGGC AGTGCCATGG ATTTCCACTG CTTCTTATGG TGGTTGGTTG 3700
GGTTTTTTGG TTTTGTTTTT TTTTTTTAAG TTTCACTCAC ATAGCCAACT 3750
CTCCCAAAGG GCACACCCCT GGGGCTGAGT CTCCAGGGCC CCCCAACTGT 3800
GGTAGCTCCA GCGATGGTGC TGCCCAGGCC TCTCGGTGCT CCATCTCCGC 3850
CTCCACACTG ACCAAGTGCT GGCCCACCCA GTCCATGCTC CAGGGTCAGG 3900
CGGAGCTGCT GAGTGACAGC TTTCCTCAAA AAGCAGAAGG AGAGTGAGTG ~~4000~~ 3950
CCTTTCCCTC CTAAAGCTGA ATCCCGGCGG AAAGCCTCTG TCCGCCTTTA ~~4050~~ 4000
CAAGGGAGAA GACAACAGAA AGAGGGACAA GAGGGTTCAC ACAGCCCAGT ~~4100~~ 4050
TCCCGTGACG AGGCTCAAAA ACTTGATCAC ATGCTTGAAT GGAGCTGGTG ~~4150~~ 4100
AGATCAACAA CACTACTTCC CTGCCGGAAT GAACTGTCCG TGAATGGTCT ~~4200~~ 4150
CTGTCAAGCG GGCCGTCTCC CTTGGCCCAG AGACGGAGTG TGGGAGTGAT ~~4250~~ 4200
TCCCAACTCC TTTCTGCAGA CGTCTGCCTT GGCATCCTCT TGAATAGGAA ~~4300~~ 4250
GATCGTTCCA CTTTCTACGC AATTGACAAA CCCGGAAGAT CAGATGCAAT ~~4350~~ 4300
TGCTCCCATC AGGGAAGAAC CCTATACTTG GTTTGCTACC CTTAGTATTT ~~4400~~ 4350
ATTACTAACC TCCCTTAAGC AGCAACAGCC TACAAAGAGA TGCTTGGAGC ~~4450~~ 4400
AATCAGAACT TCAGGTGTGA CTCTAGCAAA GCTCATCTTT CTGCCCGGCT ~~4500~~ 4450
ACATCAGCCT TCAAGAATCA GAAGAAAGCC AAGGTGCTGG ACTGTTACTG ~~4550~~ 4500
ACTTGGATCC CAAAGCAAGG AGATCATTTG GAGCTCTTGG GTCAGAGAAA ~~4600~~ 4550
ATGAGAAAGG ACAGAGCCAG CGGCTCCAAC TCCTTTCAGC CACATGCCCC ~~4650~~ 4600
AGGCTCTCGC TGCCCTGTGG ACAGGATGAG GACAGAGGGC ACATGAACAG ~~4700~~ 4650
CTTGCCAGGG ATGGGCAGCC CAACAGCACT TTTCCTCTTC TAGATGGACC ~~4750~~ 4700
CCAGCATTTA AGTGACCTTC TGATCTTGGG AAAACAGCGT CTTCCTTCTT ~~4800~~ 4750
TATCTATAGC AACTCATTGG TGGTAGCCAT CAAGCACTTC GGAATT ~~4846~~ 4796

(2) INFORMATION FOR SEQ ID NO:6
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 924
(B) TYPE: protein
(D) TOPOLOGY: linear
(ii)MOLECULE TYPE: protein
(iii) HYPOTHETICAL: no
(vi) ORIGINAL SOURCE:
(A) ORGANISM: human
(ix) FEATURE: Huntingtin-interacting protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Ser Arg Met Trp Gly His Leu Ser Glu Gly Tyr Gly Gln Leu
1 5 10 15

Cys Ser Ile Tyr Leu Lys Leu Leu Arg Thr Lys Met Glu Tyr His
20 25 30

Thr Lys Asn Pro Arg Phe Pro Gly Asn Leu Gln Met Ser Asp Arg
35 40 45

Gln Leu Asp Glu Ala Gly Glu Ser Asp Val Asn Asn Phe Phe Gln
50 55 60

Leu Thr Val Glu Met Phe Asp Tyr Leu Glu Cys Glu Leu Asn Leu
65 70 75

Phe Gln Thr Val Phe Asn Ser Leu Asp Met Ser Arg Ser Val Ser
80 85 90

Val Thr Ala Ala Gly Gln Cys Arg Leu Ala Pro Leu Ile Gln Val
95 100 105

Ile Leu Asp Cys Ser His Leu Tyr Asp Tyr Thr Val Lys Leu Leu
110 115 120

Phe Lys Leu His Ser Cys Leu Pro Ala Asp Thr Leu Gln Gly His
125 130 135

Arg Asp Arg Phe Met Glu Gln Phe Thr Lys Leu Lys Asp Leu Phe
140 145 150

Tyr Arg Ser Ser Asn Leu Gln Tyr Phe Lys Arg Leu Ile Gln Ile
155 160 165

Pro Gln Leu Pro Glu Asn Pro Pro Asn Phe Leu Arg Ala Ser Ala
170 175 180

Leu Ser Glu His Ile Ser Pro Val Val Val Ile Pro Ala Glu Ala
185 190 195

Ser Ser Pro Asp Ser Glu Pro Val Leu Glu Lys Asp Asp Leu Met
200 205 210

Asp Met Asp Ala Ser Gln Gln Asn Leu Phe Asp Asn Lys Phe Asp
215 220 225

Asp Ile Phe Gly Ser Ser Phe Ser Ser Asp Pro Phe Asn Phe Asn
230 235 240

Ser Gln Asn Gly Val Asn Lys Asp Glu Lys Asp His Leu Ile Glu
245 250 255

Arg Leu Tyr Arg Glu Ile Ser Gly Leu Lys Ala Gln Leu Glu Asn
260 265 270

Met Lys Thr Glu Ser Gln Arg Val Val Leu Gln Leu Lys Gly His
275 280 285

Val Ser Glu Leu Glu Ala Asp Leu Ala Glu Gln Gln His Leu Arg
290 295 300

Gln Gln Ala Ala Asp Asp Cys Glu Phe Leu Arg Ala Glu Leu Asp
305 310 315

Glu Leu Arg Arg Gln Arg Glu Asp Thr Glu Lys Ala Gln Arg Ser
320 325 330

Leu Ser Glu Ile Glu Arg Lys Ala Gln Ala Asn Glu Gln Arg Tyr
335 340 345

Ser Lys Leu Lys Glu Lys Tyr Ser Glu Leu Val Gln Asn His Ala
350 355 360

Asp Leu Leu Arg Lys Asn Ala Glu Val Thr Lys Gln Val Ser Met
365 370 375

Ala Arg Gln Ala Gln Val Asp Leu Glu Arg Glu Lys Lys Glu Leu
380 385 390

Glu Asp Ser Leu Glu Arg Ile Ser Asp Gln Gly Gln Arg Lys Thr
395 400 405

Gln Glu Gln Leu Glu Val Leu Glu Ser Leu Lys Gln Glu Leu Gly
410 415 420

Thr Ser Gln Arg Glu Leu Gln Val Leu Gln Gly Ser Leu Glu Thr
425 430 435

Ser Ala Gln Ser Glu Ala Asn Trp Ala Ala Glu Phe Ala Glu Leu
440 445 450

Glu Lys Glu Arg Asp Ser Leu Val Ser Gly Ala Ala His Arg Glu
455 460 465

Glu Glu Leu Ser Ala Leu Arg Lys Glu Leu Gln Asp Thr Gln Leu
470 475 480

Lys Leu Ala Ser Thr Glu Glu Ser Met Cys Gln Leu Ala Lys Asp
485 490 495

Gln Arg Lys Met Leu Leu Val Gly Ser Arg Lys Ala Ala Glu Gln
500 505 510

Val Ile Gln Asp Ala Leu Asn Gln Leu Glu Glu Pro Pro Leu Ile
515 520 525

Ser Cys Ala Gly Ser Ala Asp His Leu Leu Ser Thr Val Thr Ser
530 535 540

Ile Ser Ser Cys Ile Glu Gln Leu Glu Lys Ser Trp Ser Gln Tyr
545 550 555

Leu Ala Cys Pro Glu Asp Ile Ser Gly Leu Leu His Ser Ile Thr
560 565 570

Leu Leu Ala His Leu Thr Ser Asp Ala Ile Ala His Gly Ala Thr
575 580 585

Thr Cys Leu Arg Ala Pro Pro Glu Pro Ala Asp Ser Leu Thr Glu
590 595 600

Ala Cys Lys Gln Tyr Gly Arg Glu Thr Leu Ala Tyr Leu Ala Ser
605 610 615

Leu Glu Glu Glu Gly Ser Leu Glu Asn Ala Asp Ser Thr Ala Met
620 625 630

Arg Asn Cys Leu Ser Lys Ile Lys Ala Ile Gly Glu Glu Leu Leu
635 640 645

Pro Arg Gly Leu Asp Ile Lys Gln Glu Glu Leu Gly Asp Leu Val
650 655 660

Asp Lys Glu Met Ala Ala Thr Ser Ala Ala Ile Glu Thr Cys Thr
665 670 675

Ala Arg Ile Glu Glu Met Leu Ser Lys Ser Arg Ala Gly Asp Thr
680 685 690

Gly Val Lys Leu Glu Val Asn Glu Arg Ile Leu Arg Cys Cys Thr
695 700 705

Ser Leu Met Gln Ala Ile Gln Val Leu Ile Val Ala Ser Lys Asp
710 715 720

Leu Gln Arg Glu Ile Val Glu Ser Gly Arg Gly Thr Ala Ser Pro
725 730 735

Lys Glu Phe Tyr Ala Lys Asn Ser Arg Trp Thr Glu Gly Leu Ile
740 745 750

Ser Ala Ser Lys Ala Val Gly Trp Gly Ala Thr Val Met Val Asp
~~765~~ ~~770~~ ~~775~~
755 760 765

Ala Ala Asp Leu Val Val Gln Gly Arg Gly Lys Phe Glu Glu Leu
~~780~~ ~~785~~ ~~790~~
770 775 780

Met Val Cys Ser His Glu Ile Ala Ala Ser Thr Ala Gln Leu Val
~~795~~ ~~800~~ ~~805~~
785 790 795

Ala Ala Ser Lys Val Lys Ala Asp Lys Asp Ser Pro Asn Leu Ala
~~810~~ ~~815~~ ~~820~~
800 805 810

Gln Leu Gln Gln Ala Ser Arg Gly Val Asn Gln Ala Thr Ala Gly
~~825~~ 815 ~~830~~ 820 ~~835~~ 825

Val Val Ala Ser Thr Ile Ser Gly Lys Ser Gln Ile Glu Glu Thr
~~840~~ 830 ~~845~~ 835 ~~850~~ 840

Asp Asn Met Asp Phe Ser Ser Met Thr Leu Thr Gln Ile Lys Arg
~~855~~ 845 ~~860~~ 850 ~~865~~ 855

Gln Glu Met Asp Ser Gln Val Arg Val Leu Glu Leu Glu Asn Glu
~~870~~ 860 ~~875~~ 865 ~~880~~ 870

Leu Gln Lys Glu Arg Gln Lys Leu Gly Glu Leu Arg Lys Lys His
~~885~~ 875 ~~890~~ 880 ~~895~~ 885

Tyr Glu Leu Ala Gly Val Ala Glu Gly Trp Glu Glu Gly Thr Glu
~~900~~ 890 ~~905~~ 895 ~~910~~ 900

Ala Ser Pro Pro Thr Leu Gln Glu Val Val Thr Glu Lys Glu
~~915~~ 905 ~~920~~ 910 ~~924~~ 914

CLAIMS

1. A cDNA molecule comprising the sequence given by Seq. ID No. 1.

2. A cDNA molecule comprising the sequence given by Seq. ID No. 5.

3. A polypeptide comprising the sequence given by Seq. ID. No. 2.

4. A polypeptide comprising the sequence given by Seq. ID. No. 6.

5. A chimeric gene or plasmid comprising at least nucleotides 314 to 1955 of the Huntington's Disease gene and an activating or DNA binding domain suitable for use in a yeast multi-hybrid assay.

6. The chimeric gene or plasmid according to claim 5, wherein the Huntington's Disease gene encodes a polyglutamine tract having a length of 35 or fewer residues.

7. The chimeric gene or plasmid according to claim 5, wherein the Huntington's Disease gene encodes a polyglutamine tract having a length of 36 or more residues.

8. A method for ameliorating the effects of Huntington's disease in a patient expressing Huntingtin protein with an expanded CAG repeat region, comprising the step of increasing the amount of an expressed HD-interacting polypeptide in the brain of the patient, wherein the expressed HD-interacting polypeptide interacts less well with expanded Huntingtin than with Huntingtin having a CAG repeat region containing 15 to 35 repeats and facilitates the incorporation of Huntingtin into brain cell membranes.

9. The method according to claim 8, wherein the expressed HD-interacting polypeptide comprises the sequence given by Seq. ID No. 2.

10. An antibody which binds to a polypeptide having the sequence given by Seq. ID. No. 2.

11. The antibody of claim 10, wherein the antibody binds to amino acids 76-91 of the polypeptide having the sequence shown in Seq. ID No. 2.

12. An expression vector for expression of a gene in a mammalian host comprising a region encoding an HD-interacting polypeptide, wherein the HD-interacting polypeptide interacts less well with expanded Huntingtin than with Huntingtin having a CAG repeat region containing 15 to 35 repeats and facilitates the incorporation of Huntingtin into brain cell membranes.

13. An expression vector for expression of a gene in a mammalian host comprising a region that is the same as or complementary to Seq. ID NO. 1.

14. An expression vector for expression of a gene in a mammalian host comprising a region that is the same as or complementary to Seq. ID NO. 5.

15. The expression vector according to claims of claims 12-14, further comprising a region encoding Huntingtin having a polyglutamine tract of 35 or fewer.

16. An oligonucleotide probe having a length of from 15-40 bases which specifically and selectively hybridizes with the cDNA given by Seq. ID No. 1 or a sequence complementary thereto.

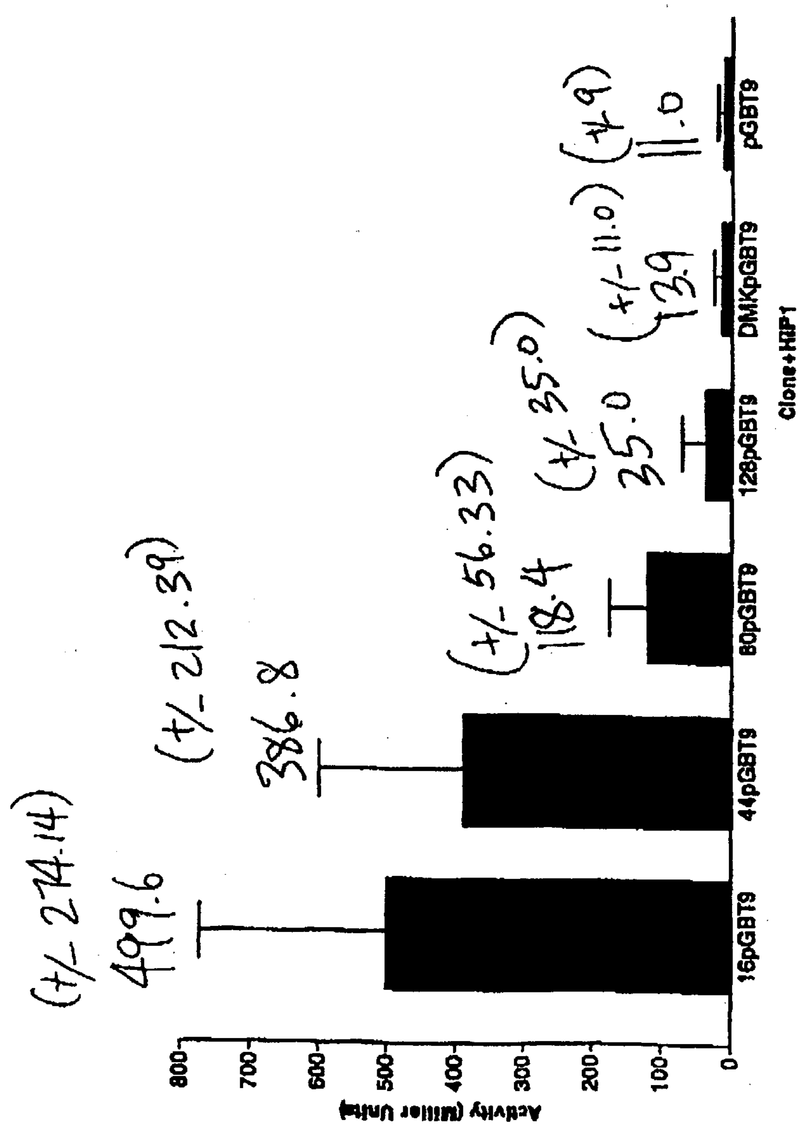
Activity (Miller Units)
800
700
600
500
400
300
200
100
0
16pGBT9
44pGBT9
80pGBT9
128pGBT9
DMKpGBT9
pGBT9
Clone+HIP1
409.6 (+/- 274.14)
386.8 (+/- 212.39)
118.4 (+/- 56.33)
35.0 (+/- 35.0)
13.9 (+/- 11.0)
11.0 (+/- 9)

Fig. 1

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/18370

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :Please See Extra Sheet.
US CL :Please See Extra Sheet.
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/320.1, 6, 69.1, 172.3; 514/44, 2; 935/62, 52, 56, 65, 34; 536/24.5, 23.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

STN, BIOSIS, MEDLINE, EMBASE, CAPLUS, WPIDS, APS, INPADOC
search terms: interacting protein, huntingtin, huntington, cag repeat,. hip, gene therapy

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	EP 0617 125 A2 (THE GENERAL HOSPITAL CORPORATION) 28 September 1994, entire document, especially pages 4-17.	8, 10-12, 16
Y	WO 94/24279 (BERGMANN ET AL.) 27 October 1994, entire document, especially pages 13 and 28-39.	8, 12 and 16
Y	EP 0 614 977 A2 (THE GENERAL HOSPITAL CORPORATION) 14 September 1994, entire document.	1-16
Y	BIAOYANG et al. Sequence of the Murine Huntington Disease Gene: Evidence for Conservation, and polymorphism in a triplet (CCG) Repeat Alternate Splicing. Human Molecular Genetics. January 1994, Vol. 3, No. 1, pages 85-92, see entire document.	8-9, 12-15

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search	Date of mailing of the international search report
15 JANUARY 1997	14 MAR 1997

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	KAREN M. HAUDA JAB for
Facsimile No. (703) 305-3230	Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/18370

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y,P	GOLDBERG et al. Cleavage of Huntingtin by apopain, a proapoptotic cysteine protease, is modulated by the polyglutamine tract. Nature Genetics. 13 August 1996, Vol. 13, No. 4, pages 442-449, see entire document.	6, 7
X,P Y,P	KALCHMAN et al. HIP-2 - A Huntingtin interacting protein: Insight into the Catabolism of the HD gene product. American Journal of Human Genetics. 02 November 1996, Vol. 59, Supplement 4, page A152, see entire document.	1-4, 12-14, and 16 5-11 and 15

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/18370

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 1-4, 9-11 and 13-16 (in part)
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

The claims recite sequence ID numbers, but no sequence disk was submitted. Due to the length of the sequences, a search could not properly be completed on the sequence ID numbers claimed.

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/18370

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (6):

A61K 38/00, 15/31, 15/09, 48/00; C12N 15/79, 15/63, 15/00; C07K 16/00; C07H 21/00

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

435/320.1, 6, 69.1, 172.3; 514/44, 2; 935/62, 52, 56, 65, 34; 536/24.5, 23.1